

## تجزیه و تحلیل بیوانفورماتیکی توالی‌های EST کلاله زعفران (*Crocus sativus* L.) به منظور تعیین جهت‌گیری کارکردی ژنوم و شبکه

### ژنی

معصومه علی اکبری<sup>۱</sup>، روح اله شاملو دشت پاگردی<sup>۲\*</sup> و اسماعیل ابراهیمی<sup>۳</sup>

تاریخ دریافت: ۹۲/۰۲/۱۹

تاریخ پذیرش: ۹۲/۰۴/۲۰

### چکیده

روش‌های ژنومیکس کارکردی نظیر تجزیه و تحلیل توالی‌های EST، امکان شناسایی، بررسی بیان و مطالعه رونوشت‌های ژنی درگیر در شبکه‌های تنظیمی و متابولیکی را فراهم آورده‌اند. در این پژوهش به منظور شناسایی جهت‌گیری کارکردی ژنوم و تعیین شبکه ژنی درگیر در تکامل کلاله زعفران، ۶۲۰۲ توالی EST مربوط به کتابخانه‌ی کلاله بالغ زعفران مورد بررسی و تحلیل قرار گرفتند. پس از پیرایش اولیه توالی‌ها، دسته بندی و هم گذاری آنها انجام شد که منتج به ایجاد ۹۱۰ unigene (۶۰۴ کانتیگ و ۳۰۴ سینگلتون) گردید. جستجوی بلاست ایکس نشان داد که ۵۷۰ unigene دارای hit مشخص در بین پروتئین‌های آرآبیدوپسیس بودند و برای سایر توالی‌ها hit مشخصی شناسایی نشد. گروه‌بندی و Gene enrichment analysis توالی‌ها آنها را در ۳۱ گروه کارکردی مختلف قرار داد که حضور ۱۲ گروه در سطح آماری ۱٪ معنی‌دار گردید. شبکه ژنی مربوط به توالی‌های با حضور بالا (بیش از ۲۰ رونوشت)، نشان داد که ارتباطات ژنی پیچیده‌ای در کلاله بالغ زعفران وجود دارد. نتایج مشخص نمود که مسیر علامت‌دهی جاسمونیک اسید و عوامل رونویسی مربوط به آن همچون *MYB21* و *Zinc finger* ها سهم بزرگی از شبکه تنظیمی را به خود اختصاص می‌دهند و نقش کلیدی‌ای در تنظیم متابولیسم اولیه و ثانویه کلاله و به خصوص متابولیسم کارتنوئیدها (به عنوان مهم‌ترین متابولیت‌های زعفران)، بر عهده دارند. ژن‌های شناسایی شده در این پژوهش می‌توانند نامزدهای مناسبی برای دست‌ورزی تکامل و متابولیسم کلاله زعفران باشند.

**واژه‌های کلیدی:** ژنومیکس کارکردی، کارتنوئید، متابولیت ثانویه.

۱، ۲ و ۳- به ترتیب دانش‌آموخته کارشناسی ارشد اصلاح نباتات، دانشجوی دکتری زراعت اصلاح نباتات و استادیار گروه زراعت و اصلاح نباتات دانشکده کشاورزی، دانشگاه شیراز.

(\*- نویسنده مسئول: Email: shamloo.r@gmail.com)

## مقدمه

زعفران<sup>۱</sup> (*Crocus sativus* L.) یا طلای سرخ<sup>۲</sup> از خانواده Iridacea، با ارزش ترین و گران‌بهارترین ادویه جهان می‌باشد. گل‌های زعفران دارای سه کلاله هستند که به‌منظور تولید ادویه زعفران جمع‌آوری و خشک می‌شوند (Ahrazem et al., 2010). در واقع کلاله‌ها مهم‌ترین بخش اقتصادی گیاه می‌باشند که به شکل‌های مختلف به-عنوان چاشنی و رنگ دهنده در تهیه غذا، در صنایع داروسازی، تولید محصولات آرایشی و مواد معطر و به‌عنوان ماده رنگ دهنده در صنایع نساجی کاربرد دارد (Srivastava et al., 2010). در سال‌های اخیر نیز اثرات ضد سرطانی و شمار دیگری از خصوصیات آنتی‌اکسیدانتی زعفران به شکل وسیعی مورد بررسی و مطالعه قرار گرفته است و به این گیاه ارزش ویژه‌ی دیگری بخشیده است (Srivastava et al., 2010).

ایران در زمینه تولید زعفران مقام اول را دارد. براساس آمارهای موجود چند صد هزار خانوار، به ویژه در شهرستان‌های قائن، تربت حیدریه، گناباد، بجستان، فردوس، ارسک، سرایان، بیرجند و کاشمر در استان‌های خراسان جنوبی و رضوی و شهرستان اقلید و شهرستان استهبان در استان فارس از این محصول روزگار می‌گذرانند (Ajam, 2007).

علی‌رغم اهمیت‌های ذکر شده برای این گیاه ارزشمند، متأسفانه آن گونه که باید و شاید در تحقیقات و طرح‌های پژوهشی راهبردی کشور به آن توجه لازم نشده است. کشوری همچون اسپانیا که وارد کننده اصلی زعفران ایران است در زمینه پژوهش‌های نوین مولکولی، بیوتکنولوژی و بیوانفورماتیکی پیرامون زعفران از سایر کشورها پیشی گرفته است.

امروزه فن‌آوری‌های امیکس<sup>۳</sup> طلایه دار روش‌های جدیدی هستند که برای کاوش در ترانسکریپتوم<sup>۴</sup>، پروتئوم<sup>۵</sup>، متابولوم<sup>۶</sup> و دیگر سطوح کارکردی ژنوم به کار می‌روند. این فن‌آوری‌ها با روشن نمودن نحوه فعالیت ژنوم در یک شرایط تکاملی یا محیطی خاص، تعیین روابط بین ژن‌ها، شناسایی نقش بخش‌های کد کننده و غیر کد کننده ژنوم و مشخص نمودن نقاط کلیدی تنظیم فرایندهای تکاملی و پاسخ به عواملی درونی و بیرونی گیاه، علاوه بر ایجاد تصویری جامع و روشن از نحوه فعالیت ژنوم، راه را برای دست‌ورزی ژنتیکی و مهندسی گیاهان در جهت بهبود صفات مورد نظر هموارتر و روشن‌تر می‌نمایند (Urano et al., 2010; Fleury et al., 2010).

روش‌های ژنومیکس کارکردی<sup>۷</sup> نظیر تولید و تجزیه تحلیل توالی‌های EST<sup>۸</sup>، تکنولوژی ریزآرایه<sup>۹</sup>، توالی‌یابی

- 1- Saffron
- 2- Red gold
- 3- Omics
- 4- Transcriptome
- 5- Proteome
- 6- Metaolome
- 7- Functional genomics
- 8- Expressed Sequenced Tags
- 9- Microarray

RNA<sup>۱</sup> یا RNA-seq امکان شناسایی، بررسی بیان و مطالعه رونوشت‌های ژنی درگیر در شبکه‌های تنظیمی و متابولیکی را فراهم آورده‌اند (Cushman & Bohnert., 2000; Ali et al., 2011).

ESTها توالی‌های کوتاه cDNA (با میانگین طول ۵۰۰ جفت باز) هستند که به وسیله توالی یابی دقیق انتهای ۵' یا ۳' همسانه‌های cDNA<sup>۲</sup> ایجاد می‌شوند. کاربرد اولیه‌ی ESTها کشف ژن است. اما از کاربردهای پیچیده‌تر آنها می‌توان به شناسایی خانواده‌های ژنی حفاظت شده بین گونه‌های مورد بررسی، کشف گوناگونی‌های ویرایش رونوشت ژن و تعیین ساختار ژن از طریق صف بندی<sup>۳</sup> و انطباق ویرایش‌ها، بررسی بیان ژن و تفسیر ژنوم<sup>۴</sup> اشاره کرد (Dong., 2005). همچنین در مواردی که توالی ژنومی یک موجود به شکل کامل مشخص نشده است، نظیر زعفران، کتابخانه‌های cDNA (و ESTهایشان) تنها رابط معتبر و جامع ژنوم باقی می‌ماند (Xoconostle-Cazares et al., 2010). داده‌های خام کتابخانه‌ها و جزئیات مربوط به هر کدام را می‌توان از پایگاه‌های داده مختلف همچون NCBI<sup>۵</sup>، TGI<sup>۶</sup>، PlandGDB<sup>۷</sup> و غیره به دست آورد. وجود این منابع غنی اطلاعات در کنار در دسترس بودن سرویس‌ها و نرم افزارهای مختلف بیوانفورماتیک راه را برای مطالعه و بررسی تغییرات ژنومی هموار نموده است (Ogata & Suzuki., 2011).

در این پژوهش به منظور شناسایی جنبه‌های مولکولی تکامل کلاله زعفران، یک کتابخانه EST مربوط به کلاله بالغ زعفران مورد تجزیه و تحلیل بیوانفورماتیکی قرار گرفت. توالی‌ها دسته بندی و نام‌گذاری شدند و پس از جستجوی بلاست، گروه‌های کارکردی ژنوم تعیین شد. همچنین رونوشت‌های با بیان بالا شناسایی و بررسی شدند و شبکه ژنی مربوط به آنها مشخص گردید.

## مواد و روش‌ها

۶۲۰۲ توالی EST مربوط به کتابخانه کلاله بالغ زعفران از پایگاه داده NCBI (<http://www.ncbi.nlm.nih.gov/>) دریافت شد. توالی‌ها به منظور شناسایی آلودگی‌های وکتوری و بررسی طول و کیفیت‌شان با استفاده از سرویس بیوانفورماتیک (EGAssembler (<http://egassembler.hgc.jp>) Masoudi-Nejad et al., 2006) مورد ارزیابی قرار گرفتند. توالی‌های مربوط به وکتور، توالی‌های تکرار شونده، توالی‌های کلروپلاستی و میتوکندریایی حذف شدند و از بین توالی‌های ارزیابی شده، توالی‌های با طول کوتاه تر از ۱۰۰ جفت باز کنار گذاشته شدند. در ادامه توالی‌های با کیفیت باقی مانده با استفاده از همان سرویس بیوانفورماتیک و با در نظر گرفتن ۸۰٪ همانندی دسته بندی<sup>۸</sup> و هم‌گذاری<sup>۹</sup> شدند. سپس جستجوی بلاست ایکس برای تمام Unigene‌ها

- 1- RNA-sequencing
- 2- Clone
- 3- Alignment
- 4- Genome annotation
- 5- <http://www.ncbi.nlm.nih.gov/>
- 6- <http://compbio.dfci.harvard.edu/tgi/>
- 7- <http://www.plantgdb.org/>
- 8- Clustering
- 9- Assembling

کانتیگ‌ها<sup>۱</sup> (توالی‌هایی با دو یا تعداد بیشتری EST ادغام شده) و سینگل‌تون‌ها<sup>۲</sup> (توالی‌های منفرد که وارد هیچ کانتیگی نشدند) با در نظر گرفتن حداکثر E-value برابر با  $10^{-5}$  در مقابل بانک توالی پروتئینی گیاه آرآبیدوپسیس به صورت Local و با نرم افزار Clcbio genomics workbench 3.6 انجام گردید. همچنین تمام کانتیگ‌ها به‌عنوان نماینده ژن‌های با بیان بالاتر مجدداً به صورت جداگانه در مقابل پایگاه nr با همان نرم افزار و E-value بلاست ایکس شدند. برای دسته‌بندی کارکردی ژن‌ها از سرویس Classify gene مربوط به موسسه فیزیولوژی گیاهی مولکولی ماکس پلانک (<http://mapman.mpimp-golm.mpg.de/general/ora/ora.shtml>) (Usadel et al., 2006) و کدهای مربوط به تمام پروتئین‌های به‌دست آمده از مرحله بلاست که دارای کمترین E-value بودند به‌عنوان ورودی سیستم استفاده گردید. همچنین، Gene enrichment analysis برای تمام کدهای پروتئینی مرحله بلاست و به‌منظور شناسایی گروه‌های کارکردی دارای حضور افزوده یا کاهش یافته معنی‌دار، با استفاده از سرویس Classify gene و با مرجع قرار دادن ژنوم آرآبیدوپسیس در سطح ۱٪ انجام شد. پس از آنالیزهای اولیه شامل هم‌گذاری، بلاست توالی‌های کانتیگ و سینگل‌تون و نهایتاً تعیین گروه‌های کارکردی و آنالیز آماری، شبکه ژنی مربوط به توالی‌های با حضور بالا که نماینده‌ای از ژن‌های با بیان بیشتر در کلاله هستند، تعیین گردید. بدین منظور کانتیگ‌های با بیش از ۲۰ توالی EST از بین سایر کانتیگ‌ها انتخاب شد. برای شناسایی این کانتیگ‌ها از برنامه نوشته شده در Python استفاده گردید. کدهای پروتئینی توالی‌های با بیان بالا که hit مشخصی در بین پروتئین‌های آرآبیدوپسیس داشتند، به‌منظور تعیین شبکه ژنی به‌عنوان ورودی سرویس GeneMANIA (<http://www.genemania.org/>) به‌کار گرفته شدند.

## نتایج و بحث

### آنالیز توالی‌های EST

پس از ارزیابی اولیه ۶۲۰۲ توالی EST، ۶۰۵۹ توالی با کیفیت و عاری از هرگونه آلودگی باقی ماند و ۱۴۳ توالی کنار گذاشته شد. میانگین طول توالی‌های باقی مانده ۳۹۲ جفت باز بود، در مجموع، این توالی‌ها ۲۳۷۵۱۲۸ جفت باز از ژنوم زعفران را پوشش داده بودند. این مقدار، چیزی در حدود ۰/۰۰۸ درصد از کل ژنوم تخمین زده شده برای زعفران یعنی ۳۰۰۰۰ میلیون جفت باز (Fiore et al. 2010) را در بر می‌گیرد. با توجه به ژنوم بسیار بزرگ زعفران و این که EST‌ها فقط نماینده بخش کد شده ژنوم می‌باشند، این مقدار کاملاً مناسب است. صاف‌بندی و هم‌گذاری توالی‌ها منتج به ایجاد ۶۰۴ توالی کانتیگ با میانگین طول ۵۰۷ جفت باز شد که در بر گیرنده ۵۷۵۳ توالی EST بودند. ۳۰۶ توالی EST نیز به‌صورت سینگل‌تون باقی ماند (جدول ۱).

در مجموع ۵۷۰ Unigene دارای hit مشخصی در بین توالی‌های پروتئینی آرآبیدوپسیس بودند و برای ۳۴۰ Unigene نیز hit مشخصی شناسایی نشد. بلاست کانتیگ‌ها در مقابل پایگاه nr هم نشان داد که ۴۱۸ توالی دارای hit مشخص و ۱۸۶ توالی بدون hit مشخص می‌باشد.

1- Cotings

2- Singletons

جدول ۱- نتایج پالایش، صف‌بندی و هم‌گذاری توالی‌های SET کلاله زعفران  
Table 1- Results of Saffron ESTs cleaning, alignment and assembling

تعداد کل توالی-های EST (Total number of EST sequences)	تعداد توالی-های با کیفیت بالا (Number of high quality ESTs)	میانگین طول توالی‌های با کیفیت بالا (جفت باز) (Average length of high quality sequences) (bp)	تعداد کانتیگ (Number of contigs)	میانگین طول توالی-های کانتیگ (جفت باز) (Average length of contigs) (bp)	تعداد سینگلتون (Number of singletons)	تعداد کل (Number of unigenes)	طول کل پوشش داده شده از ژنوم (جفت باز) (Genome coverage) (bp)
6202	6059	392	604	507	306	910	2375128

نتایج بلاست ایکس Unigeneها در برابر پروتئین‌های آرآبیدوپسیس و پایگاه nr در جدول ۲ آورده شده است.

جدول ۲- نتایج بلاست ایکس توالی‌های کانتیگ و سینگلتون در مقابل توالی‌های آرآبیدوپسیس و پایگاه nr  
Table 2- Results of Contigs and Singletons blastx against Arabidopsis and nr databases

Blastn nr (versus nr)	بلاست در مقابل پروتئین‌های آرآبیدوپسیس (Arabidopsis proteins)		Unigene
	بدون hit (Without hit)	دارای hit (With hit)	
-	137	169	سینگلتون (Singletons)
186	203	401	کانتیگ (Contigs)
-	340	570	کل

### تعیین گروه‌های کارکردی و Gene enrichment analysis

تمام Unigeneهای دارای hit در بین پروتئین‌های آرآبیدوپسیس در ۳۱ گروه کارکردی مختلف قرار گرفتند (شکل ۱). دامنه تعداد Unigeneهای اختصاص یافته به هر گروه کارکردی از یک توالی (برای گروه کارکردی تخمیر) تا ۱۰۲ توالی (برای گروه کارکردی نامعین) متغیر بود.

نتایج Gene enrichment analysis که بر مبنای آزمون کای اسکور انجام شد نشان داد که از ۳۱ گروه کارکردی شناسایی شده، ۱۲ گروه کارکردی یعنی گروه‌های گلایکولیز، گلوکونئوزنز، چرخه TCA، سنتز ATP،

متابولیسم لیپیدها، متابولیسم N، متابولیسم آمینواسیدها، متابولیسم ثانویه، پاسخ به تنش، کاهش - اکسایش، DNA و نا معین (Not-assigned)، نسبت به سایر گروه‌های کارکردی حضور افزایشی یا کاهش معنی‌داری ( $P < 0.01$ ) داشته‌اند (جدول ۳).

همان‌طور که در جدول ۳ مشخص است، به جز دو گروه کارکردی DNA و نا معین، سایر گروه‌های کارکردی نسبت حضور افزایشی داشته‌اند. بیشترین سهم افزایشی (بالای ۱۰ توالی برای هر گروه) به ترتیب مربوط به گروه -های کارکردی پاسخ به تنش، متابولیسم ثانویه، متابولیسم لیپیدها، متابولیسم آمینو اسیدها و سنتز ATP با ۲۰، ۱۷، ۱۵ و ۱۳ Unigene بود. کمترین سهم هم به گروه کارکردی متابولیسم N با ۳ Unigene اختصاص داشت.

زعفران حاوی بالغ بر ۱۵۰ ترکیب فرار و آروماتیک می‌باشد. همچنین، دارای تعداد زیادی ترکیبات فعال غیر فرار می‌باشد که بسیاری از آنها از دسته کارتئوئیدها همچون Zeaxanthin، Lycopene و انواع مختلفی از  $\alpha$ -carotene و  $\beta$ -carotene می‌باشد (Gomez-Gomez et al., 2010; Srivastava et al., 2010). اجزاء مشخصه زعفران نیز عبارتند از Crocin (مسئول تولید رنگ زعفران)، Picrocrocin (مسئول تولید طعم زعفران) و Safranal (مسئول تولید بو و عطر زعفران)، که در دسته کارتئوئیدها و کربوهیدرات قرار می‌گیرند (Gomez-Gomez et al., 2010).

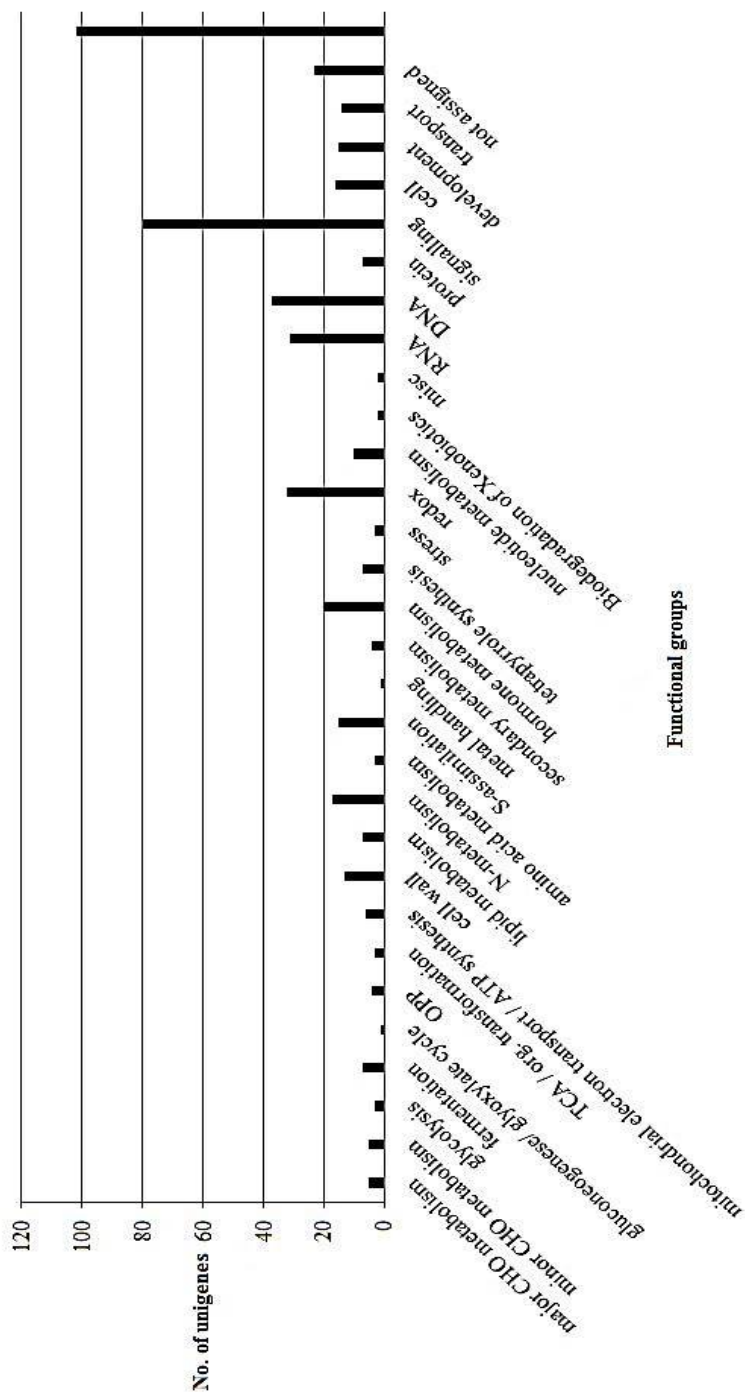
نتایج این پژوهش نشان می‌دهد که در کلاله بالغ زعفران بیشترین سهم از فعالیت ژنوم به ژن‌های پاسخ به تنش و ژن‌های درگیر در متابولیسم به خصوص متابولیسم ثانویه، لیپیدها، آمینو اسیدها و کربوهیدرات‌ها اختصاص دارد. جهت‌گیری ژنوم کلاله بالغ زعفران کاملاً با آنچه انتظار می‌رفت، یعنی جهت‌گیری برای تولید ترکیبات فرار و غیر فرار زعفران، مطابقت داشت.

#### ژن‌های با بیان بالا و شبکه ژنی مربوط به آنها

کانتینگ‌های با ۲۰ و تعداد بیشتری EST به‌عنوان ژن‌های دارای بیان بالا در کلاله زعفران در نظر گرفته شدند. بر این اساس، ۴۹ ژن شناسایی گردید که از ۲۰ تا ۵۴۵ توالی EST را در بر می‌گرفتند (شکل ۲).

۳۵ توالی از ژن‌های با بیان بالا دارای hit مشخصی بودند. نتایج تعیین شبکه ژنی نشان داد که ۳۳ ژن از این ۳۵ ژن درون شبکه ژنی قرار می‌گیرند (شکل ۳). فهرست این ژن‌ها به همراه سایر اطلاعات مربوط به آنها در جدول ۴ آورده شده است. همان‌طور که در جدول ۴ آورده شده است، ژن‌ها در ۱۲ گروه کارکردی مختلف قرار گرفتند که برخی از گروه‌های کارکردی بیش از یک ژن را به خود اختصاص دادند.

با توجه به شکل ۳، کاملاً مشخص می‌باشد که شبکه ژنی پیچیده‌ای بین ژن‌های با بیان بالا وجود دارد. البته پر واضح است که این شبکه تنها بخشی از پیچیدگی‌های کارکردی ژنوم کلاله زعفران را در برمی‌گیرد اما به دلیل حضور ژن‌هایی که بیان بسیار بالایی داشته‌اند، این بخش بسیار حائز اهمیت می‌باشد. گروه‌بندی کارکردی ژن‌های با بیان بالا نشان داد که ارتباط مستقیمی بین فعالیت گروه‌های کارکردی و بیان ژن‌ها وجود دارد چراکه بیشتر ژن‌های شبکه در گروه‌های کارکردی‌ای قرار گرفتند که خود به شکل معنی‌داری در کلاله زعفران فعال بوده‌اند (جدول ۳ و ۴).



شکل ۱- گروه‌های کارکردی مربوط به تمامی Unigene های کلابه بالغ زعفران  
 Fig. 1- Functional groups of mature Saffron stigma unigenes





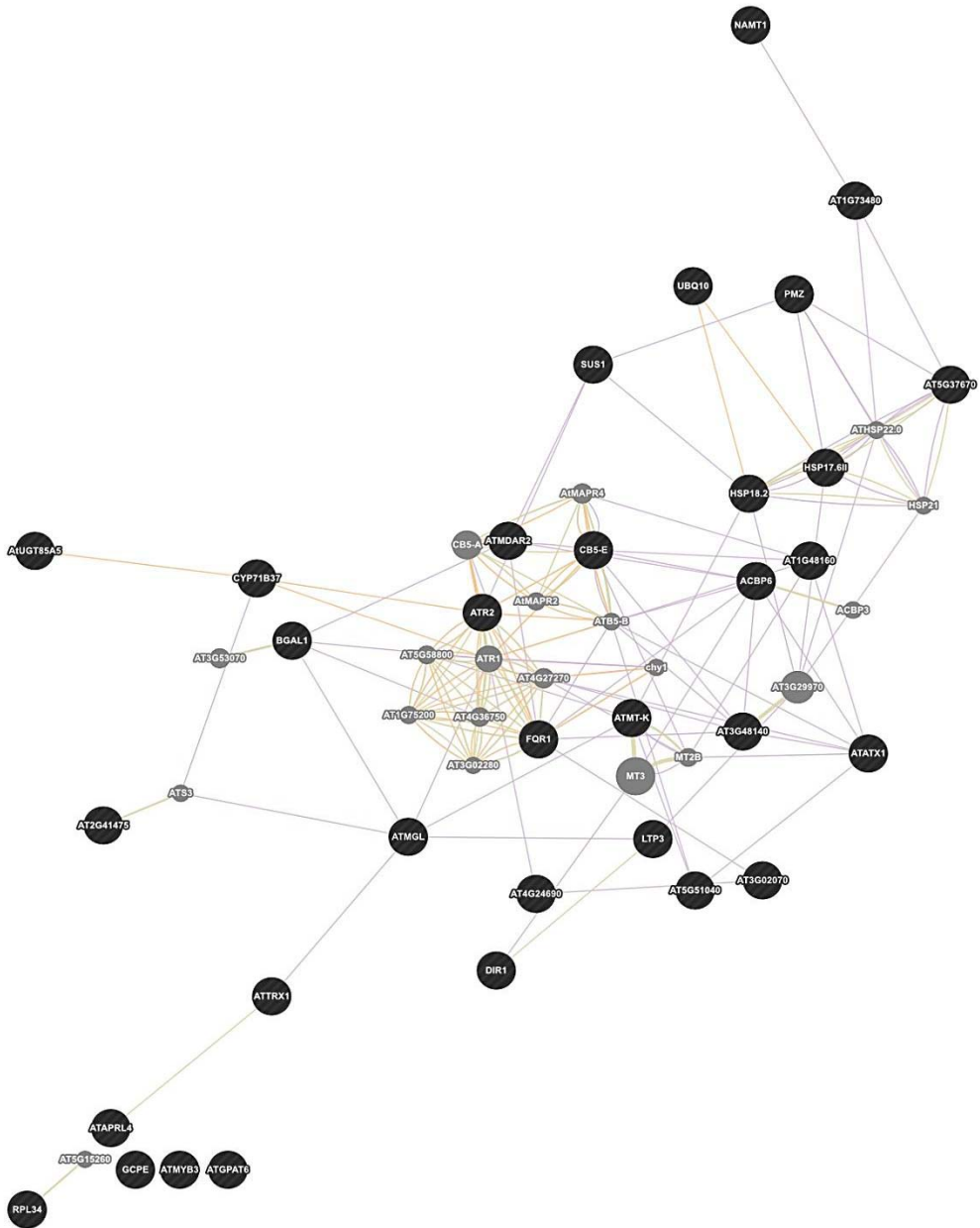
کربوهیدرات های اصلی، کدکننده آنزیمی است که در واکنش برگشت پذیر تبدیل ساکاروز و UDP<sup>1</sup> به فروکتوز و UDP - گلوکز شرکت می کند. این آنزیم نقش بسیار مهمی در متابولیسم انرژی با قرار دادن ساکاروز در مسیرهای متعدد و مرتبط با کارکردهای متابولیک، ساختاری و ذخیره ای سلول بر عهده دارد (Kladnik et al., 2004). محصولات واکنش این آنزیم متعاقباً در سایر مسیرهای متابولیک مورد استفاده قرار می گیرند چنان که UDP- گلوکز به عنوان پیش ماده در سنتز UDP- گالاکتوز، نشاسته، سلولوز و کالوز به کار گرفته می شود (Carlson et al., 2002; Kladnik et al., 2004). گزارش شده است که بیان این ژن در بافت ها و سلول های مختلف متفاوت می باشد و پروموتور این ژن کاملاً اختصاصی عمل می نماید (Kladnik et al., 2004). ژن MYB21 از دیگر ژن های مهم با بیان بالا بود. این ژن یک عامل رونویسی از زیر گروه ۱۹ خانواده R2R3 پروتئین های MYB می باشد. این عامل رونویسی یک پروتئین گل اختصاصی می باشد که در پایین دست مسیر علامت دهی جاسمونیک اسید (JA) در تکامل اندام های گل نقش بسیار کلیدی دارد (Chen et al, 2009).

به شکل جالبی یکی دیگر از عوامل رونویسی تحت تأثیر JA یعنی zinc finger AN1 domain-containing PMZ یا stress-associated protein 12 نیز در این پژوهش دارای بیان بالایی بود.

این ژن کدکننده یکی از عوامل رونویسی خانواده Zinc finger می باشد که گفته شده در تکامل اندام های گل نظیر پرچم نقش دارد (Mandaokar et al., 2006). نکته جالب دیگر این بود که در شبکه ژنی (شکل ۳) ارتباط مستقیمی بین این ژن و ژن SUS1 وجود داشت. این گونه به نظر می رسد که عوامل رونویسی مسیر علامت دهی JA در زعفران، علاوه بر نقش در تکامل اندام های گل همچون کلاله، مسیر متابولیسمی مهم کربوهیدرات ها را از طریق تنظیم بیان SUS1 تحت کنترل دارند.

در بین ژن های با بیان بالا تعدادی از پروتئین های پاسخ به تنش گرمایی یا HSP ها نیز حضور داشتند. بالا بودن فعالیت این پروتئین ها در کلاله چندان دور از انتظار نبود زیرا بسیاری از ژن های پاسخ به تنش در تکامل اندام های گیاهی نقش فعالی دارند (Reeves et al, 2012). اما نکته حائز اهمیت در مورد بالا بودن بیان این ژن ها در کلاله زعفران، ارتباط آنها با ژن های SUS1 و PMZ می باشد. در واقع به نظر می رسد که شاید بیان این گروه از ژن ها که نقش چاپرون را در سلول بازی می کنند، تحت کنترل مسیر علامت دهی JA اسید باشد.

در گروه کارکردی متابولیسم ثانویه، که در زعفران از اهمیت بالایی برخوردار می باشد، یک ژن دارای بیان بالایی بود. این ژن 4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GCPE) بود که کدکننده ی یک آنزیم کلیدی در مسیر بیوسنتز ایزوپرنوئیدها می باشد (Hunter, 2007). از آنجایی که کارتوتنوئیدها، به عنوان مهم ترین ترکیبات کلاله زعفران، خود زیر شاخه ای از ایزوپرنوئیدها می باشند (Rubio et al., 2009)، نحوه فعالیت و تنظیم مسیر بیوسنتز ایزوپرنوئیدها اهمیت ویژه ای پیدا می کند. این ژن می تواند کاندیدی برای بررسی های بیشتر در جهت شناسایی نقش آن در بیوسنتز کارتوتنوئیدهای زعفران باشد.



شکل ۳- شبکه ژنی ژن های با بیان بالا در کلالة بالغ زعفران: دایره های سیاه کدهای ژنی داده شده به سرویس و دایره های خاکستری ژن های افزوده شده به شبکه توسط سرویس می باشد.

Fig. 3- Gene network of high represented genes of mature Saffron stigma. Black circles: gene codes applied in service. Gray circles: complementary genes added from service.

جدول ۴- ژن های با بیان ( $\geq 20$  EST): کد ژنی در آرآبیدوپسیس، شرح ژن و گروه کارکردی مربوط به هر ژن آورده شده است.

**Table 4- High represented genes (EST sequences  $\geq 20$ ): Arabidopsis gene id, description and functional group of each gene is represented.**

Gene ID	Description	Functional group
AT2G41475	Embryo-specific protein 3, (ATS3)	Development
AT3G27810	transcription factor MYB21 (ATMYB3)	RNA
AT5G04370	S-adenosyl-L-methionine-dependent methyltransferase-like protein (NAMT1)	Hormone metabolism
AT1G34780	protein alfin-like 4 (ATAPRL4)	Redox
AT2G38110	glycerol-3-phosphate acyltransferase 6 (ATGPAT6)	Lipid metabolism
AT5G48485	bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin-like protein (DIR1)	Misc
AT1G22370	UDP-glucosyl transferase 85A5 (AtUGT85A5)	Misc
AT5G60600	4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GCPE)	Secondary metabolism
AT4G05320	polyubiquitin 10 (UBQ10)	Protein
AT3G02070	cysteine proteinase-like protein	Protein
AT1G66240	homolog of anti-oxidant 1 (ATATX1)	Metal handling
AT3G26330	cytochrome P450 71B37 (CYP71B37)	Misc
AT1G64660	methionine gamma-lyase (ATMGL)	Amino acid metabolism
AT3G28210	zinc finger AN1 domain-containing stress-associated protein 12 (PMZ)	RNA
AT5G59320	non-specific lipid-transfer protein 3 (LTP3)	Lipid metabolism
AT3G13750	beta galactosidase 1 (BGAL1)	Miscellaneous
AT5G51040	uncharacterized protein	Not assigned

در یک نتیجه گیری کلی از بررسی انجام شده می توان گفت در کلاله بالغ زعفران، فعالیت ژنوم به سمت افزایش متابولیسم ثانویه، متابولیسم کربوهیدرات ها، لیپیدها و آمینو اسیدها و فعال سازی مسیرهای علامت دهی و پاسخ به تحریکات که با مسیرهای تکاملی هم پوشان می باشند، گرایش پیدا می کند.

جدول ۴- (ادامه)- ژن‌های با بیان ( $\geq 20$  توالی EST): کد ژنی در آرکیدوپسیس، شرح ژن و گروه کارکردی مربوط به هر ژن آورده شده است.

**Table 4- High represented genes (EST sequences  $\geq 20$ ): Arabidopsis gene id, description and functional group of each gene is represented**

Gene ID	Description	Functional group
AT3G09390	metallothionein 2A (ATMT-K)	Metal handling
AT3G51030	thioredoxin H1 (ATTRX1)	Redox
AT3G48140	B12D protein	Development
AT5G03630	monodehydroascorbate reductase (NADH) (ATMDAR2)	Redox
AT5G20830	sucrose synthase 1 (SUS1)	Major CHO metabolism
AT5G37670	heat shock protein 15.7	Stress
AT4G24690	ubiquitin-associated (UBA) zinc-finger and PB1 domain-containing protein	Protein
AT1G73480	alpha/beta-hydrolase domain-containing protein	Not assigned
AT1G31812	acyl-CoA-binding protein 6 (ACBP6)	Lipid metabolism
AT1G48160	signal recognition particle subunit SRP19	Protein
AT1G69620	60S ribosomal protein L34-2 (RPL34)	Protein
AT5G59720	heat shock protein 18.2 (HSP18.2)	Stress
AT5G12020	heat shock protein 17.6-II (HSP17.6II)	Stress
AT5G53560	cytochrome b5 isoform 1 (CB5-E)	Misc
AT5G54500	flavodoxin-like quinone reductase 1 (FQR1)	Lipid metabolism
AT4G30210	NADPH-ferrihemoprotein reductase (ATR2)	Misc

#### منابع

- Ahrazem, O., Trapero, A., Gómez, M. D., Rubio-Moraga, A and Gómez-Gómez L. 2010. Genomic analysis and gene structure of the plant carotenoid dioxygenase 4 family: A deeper study in *Crocus sativus* L. and its allies. *Genomics* 96: 239-250.
- Ajam, M. 2007. Red gold of Iran is still unrivalled. Available at Web site <http://hamshahrionline.ir/details/38311>.
- Ali, Q., Ahsan. M., Tahir. M., Elahi. M., farooq. J. and Waseem, M. 2011. Gene expression and functional genomics approach for abiotic stress tolerance in different crop species. 2011. *IJAVMS* 5: 221-248.
- Carlson, S., J. P., Chourey<sup>1</sup>, S., Helentjaris, T. and Datta, R. 2002. Gene expression studies on developing kernels of maize sucrose synthase (SuSy) mutants show evidence for a third SuSy gene. *Plant Molecular Biology* 49: 15-29.

5. Chen, S., Li, X. Q., Zhao, A., Wang, L., Li, X., Shi, Q., Chen, M., Guo, J., Zhang, J., Qi, D. and Liu, G. 2009. Genes and Pathways Induced in Early Response to Defoliation in Rice Seedlings. *Current Issues in Molecular Biology* 11: 81-100.
6. Cushman, J. C. and Bohnert, H. J. 2000. Genomic approaches to plant stress tolerance. *Current Opinion in Plant Biology* 3: 117-124.
7. Dong, Q., Kroiss, L., Oakley, F. D., Wang, B. and V. Brendel. 2005. Comparative EST analysis in plant systems. *Methods in enzymology* 395: 400-416.
8. Fiore, A., Pizzichini, D., Diretto, G., Scossa, F. and Spano, L. 2010. Genomics and Transcriptomics of Saffron: new tools to unravel the mystery behind an attractive spice. *Functional Plant Science and Biotechnology* 4: 2-25.
9. Fleury, D., Jefferies, S., Kuchel, H. and Langridge, P. 2010. Genetic and genomic tools to improve drought tolerance in wheat. *Journal of Experimental Botany* 61: 3211-3222.
10. Gomez-Gomez, L., Rubio-Moraga, A. And Ahrazem, O. Understanding carotenoid metabolism in Saffron stigmas: Unravelling aroma and colour formation. 2010. *Functional plant science and biotechnology* 4(2): 56-63.
11. Hunter, W.N. 2007. The Non-mevalonate Pathway of Isoprenoid Precursor Biosynthesis. *The Journal of Biological Chemistry* 282 (30): 21573–21577.
12. Kladnik, A., Vilhar, B., Chourey, P.S. and Dermastia, M. 2004. Sucrose synthase isozyme SUS1 in the maize root cap is preferentially localized in the endopolyploid outer cells. *Canadian Journal of Botany* 82: 96-103.
13. Mandaokar, A., Thines, B., Shin, B., Markus Lange, B., Choi, G., Koo, Y.J., Yoo, Y. J., Choi, Y. D., Choi, G. and Browse J. 2006. Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling *The Plant Journal* 46: 984–1008.
14. Masoudi-Nejad, A., K. Tonomura, S. Kawashima, Y. Moriya, M. Suzuki, M. Itoh, M. Kanehisa, T. Endo and S. Goto. 2006. EGAssembler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. *Nucleic Acids Research* Vol. 34, Web Server Issue: 459-462.
15. Ogata, O and Suzukim H. 2011. Plant expressed sequence tags databases: practical uses and the improvement of their searches using network module analysis. *Plant Biotechnology* 28: 351-360.
16. Reeves, P. H., Ellis, C. M., Ploense, S. E., Wu, M-F., Yadav, V., Tholl, D., Che' telat, A., Haupt, I., Kennerley, B. J., Hodgens, C., Farmer, E. E., Nagpal, P. and Reed, J. W. 2012. A Regulatory Network for Coordinated Flower Maturation. *PLoS Genetics* 8(2): e1002506

17. Rubio, A., Rambla, J.L., Ahrazem, O., Granell, A. and Gomez-Gomez, L. 2009. Metabolite and target transcript analysis during *Crocus sativus* L. stigma development. *Phytochemistry* 70: 1009-1016.
18. Srivastava, R., Ahmed, H., Dharamveer, D. R. K and Saraf, S. A. 2010. *Crocus sativus* L.: A comprehensive review. *Pharmacogn Rev* 4(8): 200-208.
19. Urano, K., Kurihara, Y., Seki, M. and Shinozaki, K. 2010. 'Omics' analyses of regulatory networks in plant abiotic stress responses. *Current Opinion in Plant Biology* 13: 132-138.
20. Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., O. Blasing, E., Redestig, H., Sreenivasulu, N., Krall, L., Hannah, M. A., Poree, F., Fernie, A. R. and Stitt, M. 2006. PageMan: An interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* 7: 535.
21. Xoconostle-Cazares, B., Ramirez-Ortega, F., Flores-Elenes, L. and R. Ruiz-Medrano. 2010. Drought tolerance in crop plants. *American journal of plant physiology* 5: 241-256.

Archive of SID

## Bioinformatic Analysis of Saffron (*Crocus sativus* L.) stigma EST sequences to determining functional genome orientation and gene network

M. Aliakbari<sup>1</sup>, R. Shamloo-Dashtpajardi<sup>2\*</sup> and E. Ebrahimie<sup>3</sup>

Submitted: 09-05-2013

Accepted: 11-07-2013

### Abstract

Functional genomics methods such as Expressed Sequenced Tag (EST) analysis have provided possibilities for identification, expression analysis and study of transcripts involved in metabolic and regulatory networks. In order to identify of genome orientation and to determine gene networks involved in the evolution of saffron stigma, 6202 EST sequences from mature saffron stigma were analyzed. After initial trimming, sequences clustering and assembling resulted in 910 unigenes (604 Contigs and 304 Singleton). BLAST X revealed that 570 unigene had significant hit among the Arabidopsis protein database, whereas the remaining unigenes displayed no significant match with the any hit. Classifying and gene enrichment analysis of unigenes, put them into 31 distinct functional groups, where 12 groups of them were statistically significant at  $\alpha=0.01$ . Gene network of high represented Contigs (which had greater than 20 transcripts), showed that there is a complex gene interaction in mature saffron stigmas. Results revealed that jasmonic acid signalling pathway and its transcription factors such as MYB21 and Zinc fingers play a key role in regulating of stigma primary and secondary metabolism, especially in metabolism of carotenoids (as the most important saffron metabolites). The genes identified in this study could be good candidates for manipulating the evolution and metabolism of saffron stigma.

**Keywords:** Functional genomics, Secondary metabolite, Carotenoid

1, 2 and 3- Graduated MS of plant breeding, PhD student of plant breeding and Assistant professor of molecular genetics and genetic engineering, Faculty Agriculture, Shiraz University, Iran, respectively.

(\*- Corresponding author Email: [shamloo.r@gmail.com](mailto:shamloo.r@gmail.com))